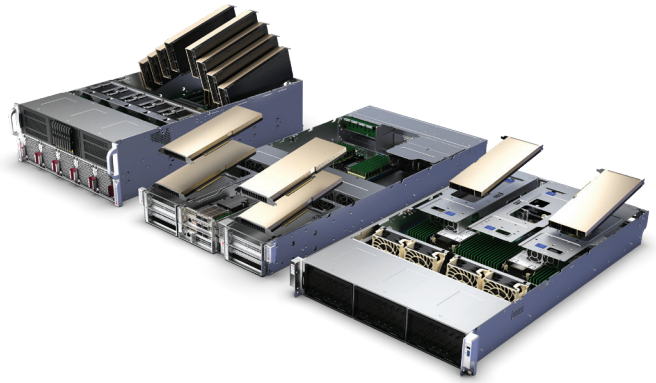
Supermicro Enterprise Solutions Featuring

NVIDIA RTX PRO™ Blackwell Server Edition GPUs

Powerful, flexible multi-workload acceleration from AI factory, to data center, to edge

Why Choose Supermicro & NVIDIA Accelerated Solutions?

Supermicro's modular and flexible accelerated solutions featuring NVIDIA RTX PRO BSE GPUs provide unprecedented performance and efficiency with industry-leading time-to-online. Available in a wide range of form factors and GPU configurations, Supermicro and NVIDIA solutions deliver up to 45x performance and 15x price/performance vs CPU-only compute, reducing infrastructure outlay and ongoing costs. With support for NVIDIA AI Enterprise software, full-stack solutions are right-sized for optimal multi-workload acceleration from AI factory, to data center, to edge.



Flexible, Modular Architectures for Industry-leading Time-to-Online

Supermicro's Building Block Solutions architectures enable faster adoption cycles for new platforms compared to OEM providers, meaning systems can be rapidly and easily adapted to meet the exact operating parameters for the deployment environment, and right-sized with the optimal number of NVIDIA RTX PRO BSE GPUs for the target workload. This accelerated development cycle combined with a large global manufacturing footprint enables rapid deployment regardless of complexity, size, or location, accelerating time-to-online and time-to-revenue for new accelerated infrastructure and turning AI potential into real-world advantage – faster.

Data Center Performance, Anywhere

Supermicro's broad system portfolio comprises a range of form factors, enabling the deployment of NVIDIA RTX PRO BSE GPUs in a wider range of enterprise computing environments, ensuring optimal performance regardless of space, power, and thermal limitations. From purpose-built AI factory SuperClusters based on NVIDIA Enterprise Reference Architectures to on-prem deployments replacing traditional CPU compute, and compact systems enabling lower latency and higher efficiency at the network edge, Supermicro and NVIDIA are able to offer right-sized solutions that deliver acceleration where it's needed most.

Complete Solutions from Supermicro and NVIDIA

Supermicro NVIDIA-Certified Systems™ guarantee compatibility with NVIDIA RTX PRO BSE GPUs, NVIDIA networking, NVIDIA AI Enterprise software, and NVIDIA Omniverse to create turnkey AI solutions. Solutions are also NVIDIA accelerated application-ready, supporting a wide range of certified third-party applications to accelerate enterprise workloads. Additionally, Supermicro offers storage solutions with leading ISVs based on NVIDIA's AI Data Platform reference architecture, which incorporate GPU acceleration into the storage platform to accelerate data vectorization, vector database searching, and inference workloads.

Do More with NVIDIA RTX PRO BSE GPUs

NVIDIA RTX PRO BSE GPUs are multi-workload accelerators designed to deliver industry-leading performance and efficiency for enterprise workloads, providing an easy upgrade path without redesigning racks, power, or cooling, harnessing Blackwell-class performance within existing environments. The NVIDIA RTX PRO 4500 BSE delivers breakthrough performance for demanding data processing, AI, video, and visual computing workloads in a single-slot, power-efficient design. Meanwhile, the NVIDIA RTX PRO™ 6000 BSE is a powerful data center GPU for AI and visual computing, accelerating demanding enterprise workloads including AI, scientific computing, graphics, and video applications.




AI Factory & Large-Scale Inference	4U 8-GPU Available in direct-connect and dual-root configurations	4U RTX PRO Server High-density NVIDIA MGX architectures	5U 8-GPU Thermally-optimized design for efficient cooling and maximum storage	3U 8-GPU Ultra-high-density configuration for on-prem AI inferencing
Model	SYS-422GA-NRT AS -4125GS-TNRT AS -4125GS-TNRT2	SYS-422GL-FNR2 SYS-422GL-NR	SYS-522GA-NRT	SYS-322GA-NR
Key Applications	AI Factory, Large-scale AI Inference, FP32 HPC	AI Factory, Large-scale AI Inference Blackwell	Large-scale AI Inference, FP32 HPC	AI Inference
GPU Support	Up to 8 NVIDIA RTX PRO 6000 Blackwell (TNRT2/NRT) Up to 4 NVIDIA RTX PRO 6000 Blackwell (TNRT)	Up to 8 NVIDIA RTX PRO 6000 Blackwell Up to 8 NVIDIA RTX PRO 4500 Blackwell	Up to 8 NVIDIA RTX PRO 6000 Blackwell Up to 8 NVIDIA RTX PRO 4500 Blackwell	Up to 4 NVIDIA RTX PRO 6000 Blackwell Up to 8 NVIDIA RTX PRO 4500 Blackwell
Networking	NVIDIA Spectrum-X Ethernet SuperNIC™ NVIDIA BlueField® DPU	NVIDIA Connect®X-8 PCIe switch board (FNR2) NVIDIA Spectrum-X Ethernet SuperNIC NVIDIA BlueField DPU (NR)	NVIDIA Spectrum-X Ethernet SuperNIC NVIDIA BlueField DPU	NVIDIA Spectrum-X Ethernet SuperNIC NVIDIA BlueField DPU
CPU	Dual Intel® Xeon® 6900 Series Dual AMD EPYC™ 9004/9005 series	Dual Intel Xeon 6900 Series	Dual Intel Xeon 6900 Series	Dual Intel Xeon 6900 Series
Storage	EDSFF E3.S 2.5" NVMe	EDSFF E1.S EDSFF E3.S	2.5" NVMe	EDSFF E1.S




Enterprise Data Center	2U RTX PRO Server Compact dual processor design for AI factories	2U MGX™ High performance DP architecture for AI and HPC	2U Hyper Flagship performance rackmount platform in UP or DP architecture	1U Hyper Flagship performance rackmount platform in DP architecture
Model	SYS-222C-TN	ARS-222GL-NR SYS-221GE-FNR2	SYS-222H-TN AS -2015HS-TN	SYS-122H-TN
Key Applications	Data Analytics, Visual Computing	Agentic AI, AI Training, Virtualization, HPC	HPC, Virtualization	HPC, Virtualization
GPU Support	Up to 2 NVIDIA RTX PRO 6000 Blackwell Up to 4 NVIDIA RTX PRO 4500 Blackwell	Up to 4 NVIDIA RTX PRO 6000 Blackwell Up to 6 NVIDIA RTX PRO 4500 Blackwell	Up to 4 NVIDIA RTX PRO 4500 Blackwell (SYS-222H-TN) Up to 2 NVIDIA RTX PRO 4500 Blackwell (AS -2105HS-TN)	1 NVIDIA RTX PRO 4500 Blackwell
Networking	NVIDIA Spectrum-X Ethernet SuperNIC NVIDIA BlueField DPU	NVIDIA Spectrum-X Ethernet SuperNIC NVIDIA BlueField DPU	NVIDIA Spectrum-X Ethernet SuperNIC NVIDIA BlueField DPU	NVIDIA Spectrum-X Ethernet SuperNIC NVIDIA BlueField DPU
CPU	Dual Intel Xeon 6700 Series	Dual NVIDIA Vera CPU (ARS-222GL-NR) Dual 5th/4th Gen Intel Xeon Scalable (SYS-221GE-FNR2)	Dual Intel Xeon 6700 Series Single AMD EPYC 9004/9005 series	Dual Intel Xeon 6700 Series
Storage	2.5" NVMe	EDSFF E1.S	2.5" NVMe	2.5" NVMe




Edge-Optimized	1U Compact Edge Server Short-depth chassis for edge data centers	2U Compact Edge Server Short-depth chassis with front or rear I/O	2U Hyper-E Flagship data center performance at the edge with front or rear I/O	2U UP RTX PRO Server Single processor architecture optimized for edge AI inference with front or rear I/O
Model	SYS-112B-FWT	AS -2116S-TNRT AS -2116S-FTNRT	SYS-222HE-TN	SYS-212GB-NR SYS-212GB-FNR
Key Applications	Edge AI, Enterprise Edge	Edge AI, Telco Edge	Edge AI, 5G Networking	Edge AI, Financial Services, Scientific Research
GPU Support	1 NVIDIA RTX PRO 4500 Blackwell	1 NVIDIA RTX PRO 4500 Blackwell	Up to 3 NVIDIA RTX PRO 4500 Blackwell	Up to 2 NVIDIA RTX PRO 6000 Blackwell Up to 4 NVIDIA RTX PRO 4500 Blackwell (FNR)
Networking	via AOC	via AOC	NVIDIA Spectrum-X Ethernet SuperNIC NVIDIA BlueField DPU	NVIDIA Spectrum-X Ethernet SuperNIC NVIDIA BlueField DPU
CPU	Single Intel Xeon 6700 Series	Single AMD EPYC 9005 series	Dual Intel Xeon 6700 Series	Single Intel Xeon 6700 Series
Storage	2.5" NVMe	2.5" NVMe/SATA	2.5" NVMe	EDSFF E1.S